Analytical Studies Branch

75 Years of Excellence / Ans d'excellence

EXTENDING HISTORICAL COMPARABILITY IN INDUSTRIAL CLASSIFICATION

by

John S. Crysdale

No. 59

Research Paper Series

Statistics Canada / Statistique Canada

Canada

ANALYTICAL STUDIES BRANCH
RESEARCH PAPER SERIES

The Analytical Studies Branch Research Paper Series provides for the circulation, on a pre-publication basis, of research conducted by Branch staff, visiting Fellows and academic associates. The Research Paper Series is intended to stimulate discussion on a variety of topics including labour, business firm dynamics, pensions, agriculture, mortality, language, immigration, statistical computing and simulation. Readers of the series are encouraged to contact the authors with comments, criticisms and suggestions. A list of titles appears inside the back cover of this paper.

Papers in the series are distributed to Statistics Canada Regional Offices, provincial statistical focal points, research institutes, and specialty libraries. Each paper is catalogued on the DOBIS computer reference system and in various Canadian university library reference systems.

To obtain a collection of abstracts of the papers in the series and/or copies of individual papers (in French or English), please contact:

Publications Review Committee
Analytical Studies Branch, Statistics Canada
24th Floor, R.H. Coats Building
Ottawa, Ontario, K1A 0T6
(613) 951-8213

EXTENDING HISTORICAL COMPARABILITY IN INDUSTRIAL CLASSIFICATION

by

John S. Crysdale

No. 59

Industry Division
Business and Trade Statistics Field
Statistics Canada
1993

Aussi disponible en français

Extending Historical Comparability in Industrial Classification

By John S. Crysdale, Statistics Canada

Weep not that the world changes--did it keep
A stable, changeless state, 'twere cause indeed to weep.
William Cullen Bryant (1824)

Abstract

The need to deal with changes in the basis of industrial classification is a perennial problem facing users of establishment-based data. A common strategy is to reclassify to a single version of the Standard Industrial Classification (SIC). This paper evaluates several automated techniques by which the statistical agency can perform that reclassification. These techniques comprise (1) using reported commodity detail, together with a set of resistance rules (2) using a one-to-one concordance and (3) using a mix of the two. Each technique is evaluated by using it to reclassify every manufacturing establishment reporting commodity detail in 1982 and by then comparing the results against the official assignments for that year. In 1982, the official series were classified and published on both a 1970 SIC and a 1980 SIC basis. The technique deemed best is the one which most closely reproduces those official assignments; it can then be used to reclassify the data of other years. The main conclusion is that a mix of commodity detail and concordance coding outperforms the alternatives, especially when used to extend classification on a 1970 SIC basis.

A non-SIC strategy, also examined here, involves finding equivalent aggregations of entire 1970 SIC industries and 1980 SIC industries, and assigning each grouping a numeric identifier. Those identifiers can then be used to recode the data of any year classified on either basis. By eliminating unusual or questionable inter-industry links from the underlying data, groupings can be kept small and homogeneous. The main disadvantage of this strategy is that the resultant industries are not as widely-recognized as those of the SIC.

Key Words: Automated industry classification, Standard Industrial Classification (SIC)

Acknowledgements

This paper originated with a reclassification request made by John Baldwin and Paul Gorecki, who, at the time, were visiting researchers in the Analytical Studies Branch. The work benefitted greatly from discussions and suggestions made by Shaila Nijhowne of Standards Division. Thanks are also due to Jack Bailey, Gerard Côté, Gil Elliot and Ken Young, all of Standards Division, Katherine Blais, Small Business and Special Surveys Division, John S. McVey, Business and Trade Statistics Field, Garnett Picot, Business and Labour Market Analysis Group and Harley Potter, formerly of Industry Division. Thanks too to George Andrusiak, Industry Division, Bruce Cooke Industry Measures and Analysis Division, Brenda Hutchinson, Industry Division, and Bruce Mitchell, National Accounts and Environment Division, for their very helpful attendance at a dry run of the presentation given at the International Conference on Establishment Surveys held in Buffalo, New York. Shortcomings are the author's.

INTRODUCTION

In Canada, the most recent change in the basis of industrial classification involved the 1983 adoption of the 1980 version of the SIC. As a result, the 171 manufacturing industries of the 1970 SIC, plus one non-manufacturing industry, converted to 236 1980 SIC manufacturing industries and three non-manufacturing industries. In many cases, the transition was simple: 79 1970 SIC industries converted on a one-to-one basis and two converted on a many-to-one basis. But, often, the transition was less simple: eleven 1970 SIC industries converted on a one-to-many basis and eighty converted on a many-to-many basis; in one case, a single many-to-many group comprised 59 1970 SIC industries and 84 1980 SIC industries.[1] [2]

The objective of this paper is to compare different ways--all fully automated--that a researcher with access to machine-readable microdata can deal with that classification break and put the data on a comparable basis. The paper deals with manufacturing establishments reporting detailed commodity data[3]; in 1982 these accounted for 58.7% of statistical units and 97.6% of manufacturing activity shipments.[4] [5]

---

[1] For further details see 'Notes on the 1980 Standard Industrial Classification in the Manufacturing Industries' in *Manufacturing Industries of Canada: national and provincial areas, 1983*, Cat. 31-203, xxiii-xcviii.

[2] For industries that are part of manufacturing on one basis and not the other, all discussion is limited to the overlap with manufacturing; mappings to industries that are not part of manufacturing on either basis are ignored.

[3] *Commodity* is used here interchangeably with product, and comprises goods of own manufacture as well as services performed on goods owned by other manufacturers (custom and repair work).

[4] Manufacturing shipments: the sum of commodity shipments, adjusted at the establishment level to net out (among other things) sales taxes and transportation charges.

There are three basic strategies used here to achieve comparable classification: (1) Extending the 1970 SIC forward in time by applying it to establishments now classified on a 1980 SIC basis. This would enable researchers to update statistical work already undertaken on a 1970 SIC basis. (2) Extending the 1980 SIC backward in time by applying it to establishments now classified on a 1970 SIC basis. This would reflect the more current model of industry structure. (3) Finding equivalent aggregations of entire 1970 and 1980 SIC industries. The resulting industries are of neither the old standard nor the new, but are closely related to each.

Three methods are employed to extend the 1970 SIC and the 1980 SIC. These involve (1) using reported product detail, along with a set of resistance rules intended to prevent establishments from flip-flopping between industries, (2) using a forced one-to-one concordance and (3) using a mix of the two.

More than one method of reclassification exists, even with full access to the microdata, due to the subjective aspects of industry classification discussed in the next section. The one-to-one concordance implicitly reflects the subjective considerations embedded in the series from which reclassification is taking place. The product detail methodology must model them explicitly.

The first section of this paper deals with the classification process followed in creating the official series. The second section discusses, in general terms, three methods of extending SIC-based classification. In the third section, those methods are evaluated by using them to reclassify manufacturing establishments reporting commodity detail in 1982 and by then comparing the results against the official assignments for that year. In 1982, data were collected on a 1970 SIC basis, but published on both bases. The main finding is that, at the 4-digit level, the industry assignments which most closely match those of the official series are achieved by bringing the 1970 SIC forward and by doing so using a mix of methods. In the fourth section, a non-SIC strategy, aggregation, is discussed. That strategy is simple to apply; its main disadvantage is that the resultant industries are not as widely recognized as those of the SIC.

---

[5] In the version of this paper published in the *Proceedings of the International Conference on Establishment Surveys*, a given establishment was not considered to report commodity detail if the questionnaire covering its activities had been completed by a related establishment and if the data for both units had been combined and if those combined data had not been reallocated by subject matter staff. For the present version of this paper, all remaining reallocation has been performed (by this author) and the corresponding establishments added to the group considered to report commodity detail. (Reallocation involved using manufacturing employment--the one item available in uncombined form for each of these establishments--to pro-rate the combined data. The results are consistent with the fact that, within each of these combinations, all constituent establishments are engaged in similar activities.) Establishments whose data were entirely estimated by the statistical agency have also been included. As a result of these extra inclusions, the percentages considered to report commodity detail and, therefore, subject to reclassification have increased from 57.0 and 96.0 to the just cited 58.7 and 97.6.

I. CLASSIFICATION IN THE OFFICIAL SERIES

Since much of this paper deals with replicating official industry assignments for manufacturing establishments reporting commodity detail, it is useful to review how those assignments are made.

Classification occurs at the 4-digit level of the SIC. Each 4-digit SIC industry is defined in terms of the manufacture of specific commodities which are said to be *primary* to that industry.[6] At the establishment level, a tentative industry assignment is calculated by grouping reported commodity outputs by primary industry and by then determining which group accounts for the largest share of commodity shipments.[7]

From the 1982 reference year to the present time, this calculation has been performed by machine.[8] The result is then compared against the establishment's existing assignment (typically last year's code; or, for births, an assignment based on *nature of business* enquiries). If the comparison indicates that the subject establishment should be considered for transfer to another industry, a print-out is produced for manual inspection. This sometimes leads to an amendment to commodity codes or shipment values. If the existing and calculated industry assignments continue to differ, a number of subjective considerations enter the process to determine whether a transfer will be immediately implemented.

One such subjective consideration involves resistance rules. Such rules are intended to prevent establishments from being transferred as a result of small shifts in output proportions, unless those shifts are seen to be permanent. The effect on industry aggregates of transfers based on small changes is disproportionate. For example, if an establishment with shipments of $100 changes industry as a result of a $1 shift in output, the sending industry will decline by 100 times that $1 shift; and the receiving industry will increase by the same factor. If the shift is only temporary, and the transfer is reversed, the impact will be felt a second time. Detailed subject matter knowledge of industry conditions and intentions will limit such transfers. There is, however, no explicit set of rules.

---

[6] Those relationships are spelled out in general terms in: *Standard Industrial Classification Manual, Revised 1970* (Cat. 12-501) Occasional, and *Standard Industrial Classification 1980* (Cat. 12-501E) Occasional. They are also spelled out in more detailed terms in commodity-to-industry concordances: the Industrial Commodity Classification (ICC) commodity to 1970 SIC industry concordance is published in *Concepts and definitions of the census of manufactures* (Cat. 31-528), Occasional, 1979; the ICC commodity to 1980 SIC industry concordance is found in Table C of *Manufacturing industries of Canada: national and provincial areas, 1983* (Cat. 31-203). During the time this paper was being written, a number of revisions were made to commodity-to-industry linkages. As a result, industry assignments for 1982 calculated now may differ from assignments calculated earlier.

[7] Shipments are used because value added cannot always be calculated at the commodity level.

[8] See Crysdale, 'Industrial Classification in the Canadian Census of Manufactures: Automated Verification Using Product Data', Discussion Paper #20, *Research Paper Series*, Analytical Studies Branch, Statistics Canada.

Another subjective consideration involves industry coverage. On occasion, an establishment may be assigned to an industry that does not account for the largest share of that establishment's output. This can happen if the establishment is such a significant part of a given industry, that its exclusion would result in serious undercoverage of the industry's activities. Such treatment is more likely to occur if the industry accounting for the largest share of the subject establishment's output is one set up to incorporate otherwise unspecified activities, and if the subject establishment cannot be artificially split between the industries involved.

Classification may also be affected by confidentiality considerations. For example, if transferring a large establishment to a small, stable industry would effectively release its confidential data, the transfer might be postponed in order to permit publication of the data for that industry.

Size significance can also be a subjective consideration. A transfer may be postponed if an establishment is judged to have an insignificant impact on industry aggregates, especially if timeliness is at risk.

In summary, the official classification of manufacturing establishments reporting commodity detail is based on a mix of objective rules and subjective considerations.

II. EXTENDING THE SIC: GENERAL DISCUSSION

In this section, the three methods of extending SIC-based classification are discussed in general terms.

Method #1: Product Detail Coding

This method involves going to the microdata and calculating an industry assignment *from scratch*. It follows closely the process used to generate the official series. There are two differences.

The first difference involves the treatment of commodities reported at a level too aggregated to be to said to be primary to just one 4-digit industry. For example, services performed on goods owned by other manufacturers (custom and repair work) are covered by insufficiently detailed classes. In 1982, too-aggregated commodities accounted for 5% of the manufacturing activity shipments of establishments reporting commodity detail.[9] In the official classification process, such activity is either made primary to the industry in which the reporting establishment is found or is made primary to *no industry*. That treatment requires that an industry assignment already exists or that manual intervention can occur. In the fully-automated approach used here, these

[9] 5.0% when reclassifying to the 1970 SIC; 5.1% when reclassifying to the 1980 SIC.

commodities are either made primary to the target classification industry to which the reporting establishment is assigned by one-to-one coding, or are made primary to the target classification industry to which the reporting establishment was assigned in the previous year.[10] [11]

The second difference involves the subjective factors discussed in the previous section. Only resistance rules are explicitly modelled here. These have been codified so that the classification process can be fully automated.

In general terms the rules used are as follows: (1) If an establishment has experienced *significant* change, it is transferred immediately. (2) Otherwise, the transfer will be made when the change is seen to be *permanent.*

As applied here, change is measured as 100 minus the following:

$$\frac{\text{Value of current year shipments primary to the industry assigned in the previous year}}{\text{Value of current year shipments primary to the industry accounting for the largest share of current year activity}} \times 100$$

This formula produces values which range from 0 to 100. The greater the value, the

---

[10] More precisely, the procedure is as follows:

Step #1: If the reporting establishment is assigned, under the originating classification, to an industry which can be forced with less than 3% error to a single class of the target classification, the commodity will be treated as primary to that target classification industry. (The 3% threshold was chosen to coincide with the threshold used by the hybrid methodology so as to have a unique characterization of each industry pair within the forced one-to-one concordance.)

Step #2: Otherwise, if the reclassification is to the 1970 SIC, the commodity will be made primary to the 1970 SIC to which the establishment as a whole was assigned in the previous year. Or, if the reclassification is to the 1980 SIC, the commodity will be made primary to the 1980 SIC industry to which the establishment as a whole is assigned in 1983.

Step #3: Otherwise, the commodity will be made primary to the target classification industry determined by Step #1, even though the originating classification industry converts with 3% or more error.

In 1982, when reclassifying to the 1970 SIC, 29.7% of establishments reporting commodity detail reported at least one such item; 30.4% when reclassifying to the 1980 SIC; these establishments accounted for 32.5% of the corresponding shipments on either basis. Too-aggregated items represented the entire output of 10.6% of establishments reporting commodity detail; 10.8% on a 1980 SIC basis; these accounted for 1.7% of the corresponding shipments on either basis. When reclassifying to the 1970 SIC, 89.9% of the commodity values were handled by Step #1, 9.4% by Step #2, and 0.7% by Step #3. When reclassifying to the 1980 SIC, the percentages were 39.5, 58.0 and 2.5.

[11] These too-aggregated commodities do not include commodities which help define industries in which a process dimension is part of the definition but is not visible in the commodity itself: (1) vertical integration in some of the pulp and paper industries, (2) joint production in the combined printing and publishing industries; these groups account for 5.6% of 1982 commodities shipped on a 1970 SIC basis, 5.5% on a 1980 SIC basis. Here, as in the official series, these are handled by first checking for the presence of selected input items (in the case of vertical integration) or selected outputs (in the case of joint production).

greater the change. Change is considered *significant* if the value produced by the formula is greater than or equal to 67. The same threshold applies for reclassification to either the 1970 SIC or the 1980 SIC. And the change (however insignificant) is considered *permanent* if the calculated industry assignment of the subject establishment remains the same for two consecutive years; in such cases, transfer occurs in that second year.[12] [13]

Table 1: Incidence of Resistance Rules, 1982
Weighted by Manufacturing Activity Shipments

Reclassification to:	1970 SIC	1980 SIC
Dominant SIC unchanged	93.5	94.9
Dominant SIC changed, test		
Delay transfer (< 67)	0.6	0.6
Transfer now (>=67)	0.5	0.8
Change persists, transfer	2.5	1.1
New, move to dominant SIC	3.0	2.6
Total	100.0	100.0

In order to link generated assignments to the official series of other years, reclassification to the 1980 SIC is performed *backward* through time; for the same reason, reclassification to the 1970 SIC is performed *forward* through time. This means that in implementing these resistance rules (and in handling too-aggregated commodities), *previous year* must be interpreted as the previous year in the reclassification process; it is not necessarily the previous calendar year.

To demonstrate the impact of this set of resistance rules, the error rates calculated later in this paper will be shown both before and after the rules are implemented. The *before* assignment is the same as is calculated by the automated edit, except for the

[12] One of the implications of this particular set of rules is that, in the reclassified data, an establishment's initial industry assignment can be carried indefinitely in the face of continual slight changes in output which do not involve the same industry in any two consecutive years. In other words, if there is a change in the industry accounting for the largest share of the establishment's shipments, but the change is not significant, the existing assignment will be maintained; then, in the following year, if a completely different industry accounts for the largest share of shipments and the change is again insignificant, the initial assignment will continue to be used.

[13] Where establishments do not have two consecutive years of commodity data, transfer is immediate.

differing treatment of too-aggregated commodities.

Method #2: Forced One-to-One Coding

This method involves reclassification from existing assignments by means of industry-level tables (see Appendices A and B) that map each 1970 SIC to just one 1980 SIC, and each 1980 SIC to just one 1970 SIC.[14] By way of example, 1970 SIC 2710 Pulp and Paper Mills, which splits into 1980 SIC 2711 Pulp Industry, 2712 Newsprint Industry, 2713 Paperboard Industry, 2714 Building Board Industry and 2719 Other Paper Industries, will be forced entirely to 1980 SIC 2712 (which accounts for the largest share of the value added of SIC 2710 in the cross-classified data of 1982[15]). All establishments assigned to 1970 SIC 2710 will be recoded to 1980 SIC 2712; none will be recoded to 1980 SIC 2711, 2713, 2714 or 2719.

Forced one-to-one coding is perhaps the simplest way of effecting 4-digit

---

[14] Each concordance was constructed by taking the 1982 official series, classified according to both the 1970 SIC and the 1980 SIC, and comparing the establishment content of each 1970 SIC industry with that of each 1980 SIC industry. The result is a list of industry pairs having establishment content in common. The list is weighted to reflect the significance of the overlap between each industry pair. Then, whenever a given originating classification industry converts to more than one target classification industry, only the pairing that accounts for the largest share of the originating classification industry is retained. Alternatives to dropping all but the most significant link for each originating classification industry are: (1) aggregation (discussed in section IV of this paper), and (2) pro-rating. The latter option involves pro-rating the data of all establishments in a given originating classification industry over all the corresponding industries of the target classification, according to (say) the shipments value of the cross-classified 1982 data. This involves splitting establishment data. Since this implies particular assumptions about input proportions, and since the reclassified data are intended for establishment-level analysis, this strategy is not pursued further. Examples of establishment-based concordances are found in Tables A and B of *Manufacturing industries of Canada: national and provincial areas, 1983.*

An alternative way of constructing a concordance is at the individual commodity level--rather than by using some bundle of commodities (such as the establishment). Such concordances may be produced by comparing the commodities primary to each 1970 SIC industry with those primary to each 1980 SIC industry. The result is a list of industry pairs having defining commodities in common. It can then be weighted to reflect actual commodity shipments; or left unweighted. An example of an unweighted commodity-level concordance is the one implied jointly by the ICC commodity to 1970 SIC industry concordance and the ICC commodity to 1980 SIC industry concordance; another is the concordance found in the *Standard Industrial Classification 1980.* The establishment-based and commodity-based concordances can yield different results: (1) An industry pair which is not present on a commodity-basis (weighted or unweighted), may occur in an establishment based concordance. This can happen if, within some establishment, the reassignment and subsequent regrouping of commodities resulting from the classification revision cause an industry not concording to the old industry to now account for the plurality of activity. (2) An industry pair which is theoretically possible on an unweighted commodity-basis may not actually be realized in the data--either because no corresponding commodity shipments have occurred (i.e., it is not present in a weighted commodity-based concordance) or because the pair did not occur as a result of a regrouping of the sort described in (1) (i.e., it is not present in an establishment-based concordance.)

[15] Where purchases of outside services have not been deducted in calculating total value added, as is the case for long-form establishments, the result is properly termed *census* total value added.

reclassification. Access to and processing of detailed product data are not required. Any researcher with a list of an industry's constituent establishments can reclassify all those establishments. In fact, reclassification need not occur at the establishment level but can occur using published aggregates. Reclassification by this method also has the merit of reflecting subjective decisions embedded in the official series. For example, it reflects the application of resistance rules--without necessitating an explicit formulation of those rules. One limitation, is that, strictly speaking, a data-based concordance applies only to the year from which it was generated (although it would not typically be used to reclassify the data of that same year). And, even in that year, its application can produce errors of inclusion and exclusion (as can the other two methods of reclassification).

Method #3: Mix of Methods

This is a mix of forced one-to-one coding and of product detail coding (with resistance rules). It takes advantage of the one-to-one mapping in reflecting subjective considerations and of the product detail approach in mirroring actual practice.

Whether product detail or one-to-one coding is used for a given originating classification industry depends on whether that industry maps well (i.e., can be forced with less than some predetermined level of error, calculated as a percentage of its own shipments total, to a single class of the target classification). If so, forcing is used. Otherwise, the product detail approach is used.

As applied here, the error threshold is 3%.[16] That level was selected after some experimentation. In mixed methods reclassification to the 1970 SIC, one-to-one coding handled 92.0% of subject shipments; for reclassification to the 1980 SIC, one-to-one coding handled 52.3% of subject shipments.

III. EXTENDING THE SIC: EMPIRICAL EVALUATION

In order to evaluate these methods, each is used to classify all establishments reporting commodity detail in 1982. Assignments are generated on both a 1970 SIC and a 1980 SIC basis. Those assignments are then compared against the official assignments of 1982, which also exist on both a 1970 SIC and a 1980 SIC basis.[17]

---

[16] Appendices A and B differentiate between industries that can be forced with less than 3% error, and those which cannot.

[17] For purposes of this paper, establishments reporting commodity detail in 1982 and classified in 1983 to 1980 SIC 3721 Chemical Fertilizer and Fertilizer Materials Industry (which was not implemented in the official series until 1983) have had the 1983 assignment made effective in the series being replicated. These account for less than 1.0% of 1982 manufacturing shipments. Establishments that were part of 1970 SIC 8930 Photographic Services, n.e.s. and became part of 1980 SIC 2821 Platemaking, Typesetting

The official assignments are treated as correct. The method which most closely replicates the official 1982 series will be deemed best. It can then be used to extend SIC classification in other years.

Error Rate Measure

The error rate measure used here will be referred to as the *percent erroneously classified*. It ranges in value from zero to one hundred, and is calculated as:

$$\frac{\text{Erroneous inclusion + Erroneous exclusion}}{\text{Official-series shipments total + Methodology-based shipments total}} \times 100$$

Erroneous inclusion is the value of shipments of establishments wrongly included in a given industry by the subject methodology; erroneous exclusion is the value wrongly excluded from that same industry.

To illustrate the calculation of this measure, consider the hypothetical case where establishments officially classified to an industry report shipments of $100 and where the subject methodology assigns establishments reporting $110 to that same industry. Also, suppose that the shipments of establishments erroneously included in this industry total $40, and that those of establishments erroneously excluded total $30. Under these circumstances, the percent erroneously classified is 33.3--i.e., (($40+$30)/($100+$110))x100.

An alternative error measure involves comparing the shipments total of the official-series industry against that of the industry generated by the subject methodology, in this case, $100 and $110, respectively. This would indicate a 10% error rate. Such a comparison of aggregates neglects the establishment content behind those totals. Consequently, it can produce misleading results. For example, if instead of generating a shipments total of $110, the subject methodology had generated a total of $100, along with $100 of erroneous inclusion and $100 of erroneous exclusion, the alternative would have indicated zero error. The alternative is not used further.

Because data users often work at the 3- and 2-digit levels of detail, the various methodologies are also assessed at those levels, using the percent erroneously classified. This involves comparing the first three (or two) digits of the 4-digit code

---

and Bindery Industry in 1983 are not present in the 1982 survey frame and, hence, are not included in the series being replicated. This group can be identified only imperfectly--by finding establishments classified in 1983 to SIC 2821 for which no 1982 data were available (this algorithm is imperfect because it could also include establishments which were brand-new in 1983); the 1983 shipments of this group correspond to less than 0.5% of 1982 manufacturing activity shipments.

generated by the subject methodology against the corresponding digits of the official 4-digit code.

Results

Table 2 shows the percent erroneously classified evaluated at the 4-, 3- and 2-digit levels, averaged on a shipments-weighted basis to the all-manufacturing level (see Appendix D for error rates averaged at the 2-digit level).

Table 2: Percent Erroneously Classified, 1982 Summarized at the All-Manufacturing Level

Reclassification to:	1970 SIC	1980 SIC
4-digit Level Evaluation		
Product Detail (no resistance)	2.8	2.8
Product Detail (with resistance)	2.5	2.3
Forced One-to-One	1.7	25.7
Mix of Methods	0.8	1.6
3-Digit Level Evaluation		
Product Detail (no resistance)	2.5	2.2
Product Detail (with resistance)	2.3	1.7
Forced One-to-One	1.1	2.9
Mix of Methods	0.8	1.1
2-Digit Level Evaluation		
Product Detail (no resistance)	1.4	1.1
Product Detail (with resistance)	1.3	0.8
Forced One-to-One	0.4	0.4
Mix of Methods	0.5	0.5

The main conclusion arising from an examination of these data is that the best results are obtained by using mixed methods. Evaluated at the 4- and 3-digit levels for reclassification to either the 1970 SIC or the 1980 SIC, the mix outperforms the other methods.

Adding a set of resistance rules to the product detail methodology lowers error rates.[18]

When evaluation occurs at higher levels of aggregation, the performance of all these methods improves. This is especially so for one-to-one coding, which improves very sharply between the 4- and 2-digit levels--indicating that most one-to-one error is internal to 3- and 2-digit industries. At the 2-digit level, one-to-one coding outperforms the mix of methods.

For reclassification to the 1980 SIC, one-to-one coding performs particularly poorly at the 4-digit level. Underlying the high error rate are 82 empty SIC classes (compared to 15 under the 1970 SIC) as well as all the erroneous inclusion to which such 100% erroneous exclusion corresponds. Those empty target classification industries exist as a result of imposing a one-to-one mapping on originating classification industries that, in fact, split.

IV. A NON-SIC STRATEGY: AGGREGATION

The two main strategies applied in this paper have involved bringing the old standard forward in time and taking the new one back. An alternative is to create a completely new classification by finding aggregations of entire 1970 SIC industries and entire 1980 SIC industries that are equivalent in terms of establishment content. The groupings are then numbered. The result is an aggregation concordance. For any establishment classified on a 1970 SIC or a 1980 SIC basis, comparably-based classification can be achieved by recoding the SIC to the new grouping number.

This strategy of *grouping up* has all the advantages listed for forced one-to-one coding. In addition, no classification error results if this concordance is used for reclassification in the year from which it was generated.

There are three disadvantages: (1) The resulting classes are not as well-known as those of the SIC. (2) There is no simple hierarchical structure. (3) There is loss of detail: the 172 classes of the 1970 SIC and the 239 classes of the 1980 SIC (referred to in the introduction) reduce to just 97 groups--one of which comprises 59 1970 SIC industries and 84 1980 SIC industries.[19]

That loss of detail derives, in part at least, because groupings are generated from

---

[18] For this reclassification method, any remaining differences from the official series are due to: (1) differences in the treatment of too-aggregated commodities, (2) differences in the treatment of subjective factors, (3) revisions to the commodity-to-industry concordances, and (4) any error in the official series.

[19] The concordance created by using the official cross-classified 1982 data and by finding *precisely* equivalent establishment content has 97 groupings. These comprise 79 one-to-one 1970 SIC to 1980 SIC conversions, 1 many-to-one conversion, 11 one-to-many conversions and 6 self-contained many-to-many groups.

actual cross-classified data. This means that unusual production behaviour or erroneous classification can result in additional industries being drawn into a given group. By excluding unusual or questionable inter-industry links in the underlying data, groupings can be prevented from growing in an unwarranted fashion. In this paper, such links are defined to be those in which the overlap between two industries accounts for less than 15% of the value added of each. By excluding those links, a much more detailed concordance has been produced. The result (see Appendix C) comprises 147 industry groupings; no SIC industry is excluded; and no grouping is unduly large. However, excluding any links means that the resulting assignments will be subject to error. That error is equal to the value of establishments whose cross-classification coincides with links deemed unusual or questionable; such error accounts for less than half of one percent of overall manufacturing activity shipments.

A similar sort of concordance is used in the Input-Output tables of the Canadian System of National Accounts.[20] The industry groupings, referred to as link-level industries or historical links, relate 1960, 1970 and 1980 SIC industries. That concordance is not a true aggregation concordance (as defined here) since the groupings do not always comprise *entire* SIC industries. In several cases, SIC industries map to more than one link-level industry. Consequently, reclassification is not always a simple recode of a given SIC industry.

CONCLUSIONS

After testing three methodologies for extending SIC-based classification, the mix of product detail and one-to-one coding was seen to outperform the other methods. It was slightly better when used to extend the 1970 SIC forward in time than when used to take the 1980 SIC back.

There are several relatively minor limitations to the extension of SIC-based classification. The first is that a number of 1970 SIC industries changed in definition while that classification was in effect. This produced breaks in the officially published series that are not a product of this reclassification.[21] These can be handled by reclassifying the underlying data to the 1982 version of the 1970 SIC. A second limitation is that the definition of manufacturing, and therefore the content of the manufacturing industries, changed with the adoption of the 1980 SIC. However, that change was only slight: less than 0.5% of the 1970 SIC version of manufacturing was dropped, and less than 0.5% of the 1980 SIC version is new. A third limitation is that the new commodity classification, an extension of the Harmonized Commodity Description and Coding System, must be linked to the 1970 SIC, before that standard can be extended beyond 1987.

---

[20] See Statistics Canada *The input-output structure of the Canadian economy, 1961-1981 (Revised data)*, Catalogue 15-510, Ottawa, 1987.

[21] See *Manufacturing Industries of Canada: national and provincial areas 1983* (Cat. 31-203), 338.

In addition, a number of changes could facilitate future exercises of this sort. First, the resistance rules used in the official series should be codified. Second, all other subjective elements, such as coverage and size significance, should also be codified. Third, a manufacturing services classification should be adopted that is sufficiently detailed to allow unique links to 4-digit industries.

An alternative strategy for achieving historical comparability, and one that is simple and highly accurate, involves the use of an aggregation concordance. By eliminating unusual or questionable inter-industry links in the underlying data, the resultant groupings are kept small and homogeneous. The main disadvantage of this strategy is that the industries are not as widely-recognized as those of the SIC.

In summary, by using a mix of methods to extend SIC-based classification, or by using the non-SIC strategy discussed here, the past twenty years of manufacturing data can be put on a comparable basis of industrial classification.

REFERENCES

[1] John S. Crysdale, 'Extending Historical Comparability in Industrial Classification' *Proceedings of the International Conference on Establishment Surveys*, (Buffalo, New York, June 1993). American Statistical Association, November 1993.

[2] John S. Crysdale, 'Industrial Classification in the Canadian Census of Manufactures: Towards Less Art and More Science' *Statistical Journal of the United Nations Economic Commission for Europe*, December 1988, Volume 5, No 4., 377-392. Also available as 'Industrial Classification in the Canadian Census of Manufactures: Automated Verification Using Product Data' *Research Paper Series*, Analytical Studies Branch, Statistics Canada, Discussion Paper #20, January 1989.

[3] Statistics Canada, *Concepts and definitions of the census of manufactures*, Catalogue 31-528, Ottawa, 1979.

[4] Statistics Canada, *The input-output structure of the Canadian economy, 1961-1981 (Revised data)*, Catalogue 15-510, Ottawa, 1987.

[5] Statistics Canada, 'Notes on the 1980 Standard Industrial Classification in the Manufacturing Industries' in *Manufacturing industries of Canada: national and provincial areas, 1983*, Catalogue 31-203, Ottawa, 1986, xxiii-xcviii.

[6] Statistics Canada, *Standard Industrial Classification Manual, Revised 1970*, Catalogue 12-501, Ottawa, 1970.

[7] Statistics Canada, *Standard Industrial Classification 1980*, Catalogue 12-501E, Ottawa, 1980.

Appendix A: Forced One-to-One Concordance, 1980 to 1970 SIC[22]

Error < 3												Error >=3	
1011	1011	1993	1899	2733	2733	3121	3160	3512	3512			1052	1050
1012	1012	1994	1899	2791	2740	3191	3150	3521	3520			1099	1089
1021	1020	1995	1832	2792	2740	3192	3150	3541	3541			1631	1650
1031	1031	2431	2431	2793	2740	3193	3150	3542	3542			1712	1740
1032	1032	2432	2431	2799	2740	3194	3150	3549	3549			1719	1792
1041	1040	2435	2432	2811	2860	3199	3150	3551	3550			1829	1810
1049	1040	2441	2441	2819	2860	3211	3210	3561	3561			1911	1852
1051	1050	2443	2441	2831	2880	3231	3230	3571	3570			1999	1899
1053	1060	2444	2441	2839	2880	3241	3241	3581	3580			2433	2431
1061	1083	2445	2442	2841	2890	3242	3243	3591	3591			2434	2431
1071	1071	2491	2392	2849	2890	3243	3242	3592	3599			2442	2441
1072	1072	2495	2460	2911	2910	3244	3242	3593	3599			2451	2450
1081	1082	2496	2480	2912	2910	3251	3250	3594	3599			2492	2431
1082	1081	2511	2511	2919	2910	3252	3250	3611	3651			2493	1750
1083	1081	2512	2513	2921	2920	3253	3250	3612	3652			2494	2310
1091	1089	2521	2520	2941	2940	3254	3250	3699	3690			2499	2441
1092	1089	2522	2520	2951	2950	3255	3250	3712	3783			2599	2599
1093	1089	2541	2543	2959	2950	3257	1880	3722	3720			2821	2870
1094	1089	2542	2544	2961	2960	3259	3250	3729	3799			3021	3010
1111	1091	2543	2541	2971	2970	3261	3260	3731	3730			3022	3010
1121	1092	2549	2541	2999	2980	3271	3270	3741	3740			3023	3020
1131	1093	2561	2560	3011	3010	3281	3280	3751	3750			3053	3051
1141	1094	2581	2580	3029	3020	3299	3290	3761	3760			3256	1650
1211	1510	2591	2591	3031	3031	3311	3310	3771	3770			3372	3360
1221	1530	2592	2593	3032	3039	3321	3320	3791	3791			3562	3562
1511	1623	2593	2593	3039	3039	3331	3330	3792	3799			3599	3530
1521	1629	2611	2619	3041	3041	3332	2680	3799	3799			3711	3782
1599	1629	2612	2619	3042	3042	3333	3399	3913	3912			3721	3782
1611	1650	2619	2619	3049	3042	3341	3340	3914	3914			3911	3911
1621	1650	2641	2640	3051	3059	3351	3350	3921	3920			3912	3911
1691	2733	2649	2640	3052	3059	3352	3350	3922	3920			3999	3999
1699	1650	2691	2660	3059	3059	3359	3350	3931	3931				
1711	1720	2692	2660	3061	3060	3361	3180	3932	3932				
1713	1799	2699	2660	3062	3060	3362	3180	3971	3970				
1811	1831	2711	2710	3063	3060	3369	3180	3991	3991				
1821	1820	2712	2710	3069	3060	3371	3360	3992	3992				
1831	2391	2713	2710	3071	3070	3379	3360	3993	3993				
1921	1860	2714	2710	3081	3080	3381	3380	3994	3994				
1931	1872	2719	2710	3091	3090	3391	3391	6012	1072				
1991	1892	2721	2720	3092	3090	3392	3399	6213	2611				
1992	1894	2731	2731	3099	3090	3399	3399	9213	1072				
		2732	2732	3111	3110	3511	3511						

[22] SIC names are found in *Standard Industrial Classification Manual, Revised 1970* and in *Standard Industrial Classification, 1980*. The relationships shown here are consistent with those of the full concordance appearing in *Manufacturing industries of Canada: national and provincial areas, 1983*, Cat. 31-203. Both concordances are based on the data for all records, not just establishments reporting commodity detail. In order to have general applicability, both concordances are based on a combination of 1982 and 1983 data. The 1983 data are limited to: (1) identifying establishments reporting in 1982 are classified to SIC 3721 in 1983 so that that assignment can be made effective in 1982; (2) identifying establishments added to manufacturing in 1983 from 1970 SIC 8930, and including these (see footnote 16). This table is divided into two error groups. That error is equal to the proportion of the originating classification industry that, according to the cross-classified data from which the concordance was generated, properly belongs to industries other than the single target classification industry to which the subject industry is forced. Imposing a one-to-one relationship results in the complete exclusion of sixteen target classification industries belonging to the relevant set (including one non-manufacturing industry); these are 1970 SIC: 1624, 1840, 1851, 1871, 1891, 1893, 2491, 2492, 2499, 2592, 3781, 3913, 3915, 3996, 3998, 8930.

Appendix B: Forced One-to-One Concordance, 1970 to 1980 SIC[23]

Error < 3		1880	3257	2970	2971	3570	3571	Error >=3		3020	3029
1011	1011	1891	1999	2980	2999	3580	3581	1040	1049	3031	3031
1012	1012	1892	1991	3041	3041	3591	3591	1050	1052	3039	3039
1020	1021	1894	1992	3051	3053	3651	3611	1072	1072	3042	3049
1031	1031	2310	2494	3080	3081	3652	3612	1081	1083	3059	3059
1032	1032	2391	1831	3110	3111	3690	3699	1089	1099	3060	3062
1060	1053	2432	2435	3160	3121	3720	3722	1629	1599	3070	3071
1071	1071	2442	2445	3210	3211	3730	3731	1650	1699	3090	3099
1082	1081	2450	2451	3230	3231	3740	3741	1799	1713	3150	3199
1083	1061	2460	2495	3241	3241	3750	3751	1832	1829	3180	3361
1091	1111	2480	2496	3243	3242	3760	3761	1893	1999	3242	3243
1092	1121	2491	2493	3260	3261	3770	3771	1899	1994	3250	3251
1093	1131	2492	2499	3270	3271	3781	3711	2392	2491	3350	3351
1094	1141	2511	2511	3280	3281	3791	3791	2431	2433	3360	3379
1510	1211	2513	2512	3290	3299	3912	3913	2441	2442	3399	3333
1530	1221	2543	2541	3310	3311	3913	3999	2499	2499	3599	3594
1623	1511	2544	2542	3320	3321	3914	3914	2520	2522	3782	3711
1624	1712	2560	2561	3330	3331	3915	3999	2541	2543	3783	3712
1720	1711	2580	2581	3340	3341	3931	3931	2593	2592	3799	3799
1740	1712	2591	2591	3380	3381	3932	3932	2619	2611	3911	3911
1750	2493	2592	2599	3391	3391	3970	3971	2640	2641	3920	3921
1792	1719	2599	2599	3511	3511	3991	3991	2660	2692		
1810	1829	2611	6213	3512	3512	3992	3992	2710	2712		
1820	1821	2680	3332	3520	3521	3993	3993	2733	1691		
1831	1811	2720	2721	3530	3599	3994	3994	2740	2799		
1840	1999	2731	2731	3541	3541	3996	3999	2860	2819		
1851	1911	2732	2732	3542	3542	3998	3999	2880	2839		
1852	1911	2870	2821	3549	3549	3999	3999	2890	2841		
1860	1921	2920	2921	3550	3551	8930	2821	2910	2919		
1871	1999	2940	2941	3561	3561			2950	2959		
1872	1931	2960	2961	3562	3562			3010	3011		

[23] SIC names are found in *Standard Industrial Classification Manual, Revised 1970* and in *Standard Industrial Classification, 1980.* The relationships shown here are consistent with those of the full concordance appearing in *Manufacturing industries of Canada: national and provincial areas, 1983*, Cat. 31-203. Both concordances are based on the data for all records, not just establishments reporting commodity detail. In order to have general applicability, both concordances are based on a combination of 1982 and 1983 data. The 1983 data are limited to: (1) identifying establishments reporting in 1982 are classified to SIC 3721 in 1983 so that that assignment can be made effective in 1982; (2) identifying establishments added to manufacturing in 1983 from 1970 SIC 8930, and including these (see footnote 16). This table is divided into two error groups. That error is equal to the proportion of the originating classification industry that, according to the cross-classified data from which the concordance was generated, properly belongs to industries other than the single target classification industry to which the subject industry is forced. Imposing a one-to-one relationship results in the complete exclusion of 82 target classification industries belonging to the relevant set (including two non-manufacturing industries); these are 1980 SIC: 1041, 1051, 1082, 1091, 1092, 1093, 1094, 1521, 1611, 1621, 1631, 1993, 1995, 2431, 2432, 2434, 2441, 2443, 2444, 2492, 2521, 2549, 2593, 2612, 2619, 2649, 2691, 2699, 2711, 2713, 2714, 2719, 2733, 2791, 2792, 2793, 2811, 2831, 2849, 2911, 2912, 2951, 3021, 3022, 3023, 3032, 3042, 3051, 3052, 3061, 3063, 3069, 3091, 3092, 3191, 3192, 3193, 3194, 3244, 3252, 3253, 3254, 3255, 3256, 3259, 3352, 3359, 3362, 3369, 3371, 3372, 3392, 3399, 3592, 3593, 3721, 3729, 3792, 3912, 3922, 6012, 9213.

Appendix C: Aggregation Concordance, Excluding Selected Links[24] [25]

ID#	SIC70	SIC80
001	1011	1011
002	1012	1012
003	1020	1021
004	1031	1031
005	1032	1032
006	1040	1041 1049
007	1050 1089	1051 1052 1091 1092 1093 1094 1099
008	1060	1053
009	1071	1071
010	1072	1072 6012 9213
011	1081	1082 1083
012	1082	1081
013	1083	1061
014	1091	1111
015	1092	1121
016	1093	1131
017	1094	1141
018	1510	1211
019	1530	1221
020	1623	1511
021	1624 1740	1712
022	1629	1521 1599
023	1650 3250	1611 1621 1631 1699 3251 3252 3253 3254 3255 3256 3259
024	1720	1711
025	1750 2491	2493
026	1792 1799	1713 1719
027	1810 1832	1829 1995

[24] Abbreviations: ID#=numeric identifier.

[25] SIC names are found in *Standard Industrial Classification Manual, Revised 1970* and in *Standard Industrial Classification, 1980*. The relationships shown here are consistent with those of the full concordance appearing in *Manufacturing industries of Canada: national and provincial areas, 1983*, Cat. 31-203. Both concordances are based on the data for all records, not just establishments reporting commodity detail. In order to have general applicability, both concordances are based on a combination of 1982 and 1983 data. The 1983 data are limited to: (1) identifying establishments reporting in 1982 are classified to SIC 3721 in 1983 so that that assignment can be made effective in 1982; (2) identifying establishments added to manufacturing in 1983 from 1970 SIC 8930, and including these (see footnote 16). For each of the following 83 pairs, the overlap between the 1970 SIC industry and the 1980 SIC industry accounts for less than 15% of the value added of the 1970 SIC industry and less than 15% of the value added of the 1980 SIC industry; these SIC70-SIC80 pairs are therefore deemed unusual or questionable links, and have been excluded from the underlying data used to construct this concordance: 1081-1093 1081-1099 1089-1072 1629-1699 1650-1691 1799-1999 1831-1829 1832-1811 1893-2445 1894-1829 1894-2434 1899-1821 1899-1829 1899-1911 2392-1999 2392-2432 2392-2442 2392-2443 2392-2451 2392-2493 2392-2494 2392-2499 2431-2442 2441-1712 2441-2433 2441-2445 2441-2451 2441-2492 2441-2495 2480-2499 2499-1999 2499-2433 2499-2493 2513-0412 2619-2699 2660-1699 2660-2611 2660-2619 2660-2641 2660-2649 2733-1631 2860-2619 2860-2821 2910-3099 3010-3042 3020-3021 3031-3562 3039-3023 3042-1719 3042-3022 3042-3071 3042-3091 3042-3099 3059-3042 3059-3053 3070-3121 3070-3199 3090-3042 3090-3049 3090-3053 3090-3071 3090-3911 3090-3931 3150-3069 3150-3111 3150-3241 3150-3359 3150-3799 3180-3352 3180-3372 3250-3391 3290-3111 3310-3191 3350-3399 3350-3911 3360-3911 3399-3372 3399-3379 3399-3912 3652-3799 3781-3712 3799-3791 3913-3912.

Appendix C (concluded)

ID#	SIC70	SIC80
028	1820	1821
029	1831	1811
030	1840 1871 1891 1893 1899	1993 1994 1999
031	1851 1852	1911
032	1860	1921
033	1872	1931
034	1880	3257
035	1892	1991
036	1894	1992
037	2310	2494
038	2391	1831
039	2392 2431 2441 2492 2499	2431 2432 2433 2434 2441 2442 2443 2444 2491 2492 2499
040	2432	2435
041	2442	2445
042	2450	2451
043	2460	2495
044	2480	2496
045	2511	2511
046	2513	2512
047	2520	2521 2522
048	2541	2543 2549
049	2543	2541
050	2544	2542
051	2560	2561
052	2580	2581
053	2591	2591
054	2592 2599	2599
055	2593	2592 2593
056	2611	6213
057	2619	2611 2612 2619
058	2640	2641 2649
059	2660	2691 2692 2699
060	2680	3332
061	2710	2711 2712 2713 2714 2719
062	2720	2721
063	2731	2731
064	2732	2732
065	2733	1691 2733
066	2740	2791 2792 2793 2799
067	2860	2811 2819
068	2870 8930	2821
069	2880	2831 2839
070	2890	2841 2849
071	2910	2911 2912 2919
072	2920	2921
073	2940	2941
074	2950	2951 2959
075	2960	2961
076	2970	2971
077	2980	2999
078	3010 3020	3011 3021 3022 3023 3029
079	3031	3031
080	3039	3032 3039
081	3041	3041
082	3042	3042 3049
083	3051	3053
084	3059	3051 3052 3059
085	3060	3061 3062 3063 3069
086	3070	3071
087	3080	3081
088	3090	3091 3092 3099
089	3110	3111

Appendix C (concluded)

ID#	SIC70	SIC80
090	3150	3191 3192 3193 3194 3199
091	3160	3121
092	3180	3361 3362 3369
093	3210	3211
094	3230	3231
095	3241	3241
096	3242	3243 3244
097	3243	3242
098	3260	3261
099	3270	3271
100	3280	3281
101	3290	3299
102	3310	3311
103	3320	3321
104	3330	3331
105	3340	3341
106	3350	3351 3352 3359
107	3360	3371 3372 3379
108	3380	3381
109	3391	3391
110	3399	3333 3392 3399
111	3511	3511
112	3512	3512
113	3520	3521
114	3530 3599	3592 3593 3594 3599
115	3541	3541
116	3542	3542
117	3549	3549
118	3550	3551
119	3561	3561
120	3562	3562
121	3570	3571
122	3580	3581
123	3591	3591
124	3651	3611
125	3652	3612
126	3690	3699
127	3720	3722
128	3730	3731
129	3740	3741
130	3750	3751
131	3760	3761
132	3770	3771
133	3781 3782 3783	3711 3712 3721
134	3791	3791
135	3799	3729 3792 3799
136	3911	3911 3912
137	3912	3913
138	3913 3915 3996 3998 3999	3999
139	3914	3914
140	3920	3921 3922
141	3931	3931
142	3932	3932
143	3970	3971
144	3991	3991
145	3992	3992
146	3993	3993
147	3994	3994

Appendix D: Percent Erroneously Classified, Summarized by 2-Digit SIC[26]

	4-digit Evaluation				3-digit Evaluation				2-digit Evaluation			
	Nres	Res	Con	Mix	Nres	Res	Con	Mix	Nres	Res	Con	Mix
1970 SIC												
10 Food & Bev	0.6	0.7	0.7	0.3	0.3	0.4	0.4	0.3	0.0	0.0	0.0	0.0
15 Tobacco	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
16 Rubr/Plstc	5.1	5.4	4.1	2.2	5.2	5.4	3.2	2.2	5.0	5.2	3.2	2.2
17 Leather	1.5	1.5	6.1	1.3	1.5	1.5	4.4	1.2	1.3	1.3	4.6	1.1
18 Textile	3.6	3.3	21.2	2.4	2.7	2.6	14.1	2.0	1.2	1.1	0.1	0.6
23 Knitting	18.1	17.4	16.7	16.0	18.8	17.7	14.4	15.6	19.7	17.9	12.4	15.6
24 Clothing	7.4	8.9	10.7	7.5	7.4	8.8	10.7	7.4	6.2	5.5	3.0	4.6
25 Wood	4.2	3.4	0.8	0.4	4.4	3.4	0.0	0.1	0.6	0.4	0.0	0.0
26 Furniture	4.8	4.8	0.3	0.3	4.8	4.8	0.3	0.3	1.4	1.1	0.0	0.0
27 Paper	2.0	2.0	0.3	0.1	2.0	1.9	0.3	0.1	0.8	0.7	0.3	0.1
28 Print/Publ	3.0	2.9	0.2	0.7	3.0	2.9	0.2	0.7	0.9	0.9	0.0	0.2
29 Prim Metal	2.4	2.3	0.0	0.1	2.4	2.3	0.0	0.1	1.2	1.3	0.0	0.1
30 Metal Fab	8.3	6.8	1.8	1.8	7.3	6.5	1.5	1.6	3.6	3.2	0.5	0.4
31 Machinery	3.3	2.8	0.3	0.3	3.3	2.8	0.3	0.3	3.3	2.8	0.3	0.3
32 Transp Eqp	1.5	1.4	0.3	0.3	1.3	1.2	0.3	0.3	1.1	1.1	0.3	0.3
33 Electrical	4.8	3.8	0.8	1.3	4.8	3.9	0.8	1.3	2.6	1.9	0.2	0.4
35 Non-met Min	3.7	3.3	1.7	0.6	1.8	1.3	1.7	0.6	1.5	1.1	0.4	0.5
36 Petro/Coal	0.6	0.6	0.0	0.0	0.6	0.6	0.0	0.0	0.1	0.1	0.0	0.0
37 Chemical	2.8	2.5	3.0	0.8	2.5	2.3	0.0	0.5	0.8	0.6	0.0	0.5
39 Misc Mfg	8.7	7.0	6.2	4.2	7.5	5.6	2.5	3.7	7.1	5.3	0.9	3.4
99 Total	2.8	2.5	1.7	0.8	2.5	2.3	1.1	0.8	1.4	1.3	0.4	0.5
1980 SIC												
10 Food	1.0	0.9	22.7	0.5	0.9	0.8	1.0	0.5	0.2	0.2	0.2	0.2
11 Beverage	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
12 Tobacco	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
15 Rubber	2.4	2.0	14.5	2.0	2.4	2.0	14.5	2.0	2.3	1.9	0.0	1.9
16 Plastic	7.7	6.9	58.4	6.3	7.7	6.9	58.3	6.3	4.3	3.5	7.6	2.9
17 Leather	0.8	0.6	1.9	0.5	0.6	0.4	0.1	0.3	0.6	0.4	0.1	0.3
18 Prim Txtl	2.3	0.5	1.8	0.3	2.3	0.5	1.8	0.3	1.0	0.3	1.7	0.2
19 Textile	2.3	1.7	35.6	1.1	1.9	1.3	2.5	0.9	1.6	1.1	2.5	0.9
24 Clothing	4.1	1.7	56.8	1.7	2.8	1.2	14.8	1.1	0.5	0.2	0.1	0.2
25 Wood	1.8	1.5	11.5	0.6	1.4	1.2	0.0	0.3	0.6	0.5	0.0	0.3
26 Furniture	6.2	5.4	62.8	5.3	3.4	2.4	0.3	2.3	1.4	1.2	0.0	1.1
27 Paper	2.4	2.3	60.7	2.0	2.1	1.9	0.8	1.7	0.6	0.5	0.8	0.5
28 Print/Publ	3.6	3.6	24.0	3.4	2.8	2.7	0.2	2.5	1.0	0.9	0.0	0.7
29 Prim Metal	2.5	1.7	24.0	0.4	2.4	1.6	0.0	0.2	1.2	0.6	0.0	0.2
30 Fab Metal	8.1	7.6	43.2	6.2	6.5	5.9	4.1	4.5	3.4	3.0	0.5	1.9
31 Machinery	6.4	6.0	58.8	4.2	3.6	3.3	0.3	1.4	3.5	3.2	0.3	1.4
32 Transp Eqp	1.9	1.5	19.8	0.9	1.1	0.8	0.4	0.4	0.9	0.6	0.4	0.4
33 Electrical	3.8	3.5	46.0	2.7	3.2	3.0	5.0	2.2	1.7	1.5	0.1	0.9
35 Non-met Min	3.6	3.2	10.5	0.9	1.5	1.1	0.4	0.9	1.2	1.0	0.4	0.9
36 Petro/Coal	0.6	0.4	0.0	0.0	0.6	0.4	0.0	0.0	0.1	0.0	0.0	0.0
37 Chemical	3.6	2.3	11.6	1.6	3.6	2.2	10.0	1.5	0.6	0.3	0.0	0.2
39 Other Mfg	6.6	5.4	30.3	2.5	4.8	3.5	0.9	1.4	4.1	3.0	0.9	1.1
99 Total	2.8	2.3	25.7	1.6	2.2	1.7	2.9	1.1	1.1	0.8	0.4	0.5

[26] **Abbreviations**: Nres=product detail coding (no resistance rules), Res=product detail coding (with resistance rules), Con=one-to-one concordance, Mix=mix of methods. In this table, there is no differentiation between true zero and zero produced by rounding.

ANALYTICAL STUDIES BRANCH
RESEARCH PAPER SERIES

No.

1. *Behavioural Response in the Context of Socio-Economic Microanalytic Simulation,* ***Lars Osberg***

2. *Unemployment and Training,* ***Garnett Picot***

3. *Homemaker Pensions and Lifetime Redistribution,* ***Michael Wolfson***

4. *Modelling the Lifetime Employment Patterns of Canadians,* ***Garnett Picot***

5. *Job Loss and Labour Market Adjustment in the Canadian Economy,* ***Garnett Picot and Ted Wannell***

6. *A System of Health Statistics: Toward a New Conceptual Framework for Integrating Health Data,* ***Michael C. Wolfson***

7. *A Prototype Micro-Macro Link for the Canadian Household Sector,* ***Hans J. Adler and Michael C. Wolfson***

8. *Notes on Corporate Concentration and Canada's Income Tax,* ***Michael C. Wolfson***

9. *The Expanding Middle: Some Canadian Evidence on the Deskilling Debate,* ***John Myles***

10. *The Rise of the Conglomerate Economy,* ***Jorge Niosi***

11. *Energy Analysis of canadian External Trade: 1971 and 1976,* ***K.E. Hamilton***

12. *Net and Gross Rates of Land Concentration,* ***Ray D. Bollman and Philip Ehrensaft***

13. *Cause-Deleted Life Tables for Canada (1972 to 1981): An Approach Towards Analyzing Epidemiologic Transition,* ***Dhruva Nagnur and Michael Nagrodski***

14. *The Distribution of the Frequency of Occurence of Nucleotide Subsequences, Based on Their Overlap Capability,* ***Jane F. Gentleman and Ronald C. Mullin***

15. *Immigration and the Ethnolinguistic Character of Canada and Quebec,* ***Réjean Lachapelle***

16. *Integration of Canadian Farm and Off-Farm Markets and the Off-Farm Work of Women, Men and Children,* ***Ray D. Bollman and Pamela Smith***

17. *Wages and Jobs in the 1980s: Changing Youth Wages and the Declining Middle,* ***J. Myles, G. Picot and T. Wannell***

18. *A Profile of Farmers with Computers,* ***Ray D. Bollman***

19. *Mortality Risk Distributions: A Life Table Analysis,* ***Geoff Rowe***

20. *Industrial Classification in the Canadian Census of Manufactures: Automated Verification Using Product Data,* ***John S. Crysdale***

21. *Consumption, Income and Retirement,* ***A.L. Robb and J.B. Burbridge***

22. *Job Turnover in Canada's Manufacturing Sector,* ***John R. Baldwin and Paul K. Gorecki***

23. *Series on The Dynamics of the Competitive Process,* ***John R. Baldwin and Paul K. Gorecki***

    A. *Firm Entry and Exit Within the Canadian Manufacturing Sector.*
    B. *Intra-Industry Mobility in the Canadian Manufacturing Sector.*
    C. *Measuring Entry and Exit in Canadian Manufacturing: Methodology.*
    D. *The Contribution of the Competitive Process to Productivity Growth: The Role of Firm and Plant Turnover.*
    E. *Mergers and the Competitive Process.*
    F. *(in preparation)*
    G. *Concentration Statistics as Predictors of the Intensity of Competition.*
    H. *The Relationship Between Mobility and Concentration for the Canadian Manufacturing Sector.*

24. *Mainframe SAS Enhancements in Support of Exploratory Data Analysis,* ***Richard Johnson and Jane F. Gentleman***

25. *Dimensions of Labour Market Change in Canada: Intersectoral Shifts, Job and Worker Turnover,* ***John R. Baldwin and Paul K. Gorecki***

26. *The Persistent Gap: Exploring the Earnings Differential Between Recent Male and Female Postsecondary Graduates,* ***Ted Wannell***

27. *Estimating Agricultural Soil Erosion Losses From Census of Agriculture Crop Coverage Data,* ***Douglas F. Trant***

28. *Good Jobs/Bad Jobs and the Declining Middle: 1967-1986,* ***Garnett Picot, John Myles, Ted Wannell***

29. *Longitudinal Career Data for Selected Cohorts of Men and Women in the Public Service, 1978-1987,* ***Garnett Picot and Ted Wannell***

30. *Earnings and Death - Effects Over a Quarter Century,* ***Michael Wolfson, Geoff Rowe, Jane F. Gentleman adn Monica Tomiak***

31. *Firm Response to Price Uncertainty: Tripartite Stabilization and the Western Canadian Cattle Industry,* ***Theodore M. Horbulyk***

32. *Smoothing Procedures for Simulated Longitudinal Microdata,* ***Jane F. Gentleman, Dale Robertson and Monica Tomiak***

33. *Patterns of Canadian Foreign Direct Investment Abroad,* ***Paul K. Gorecki***

34. *POHEM - A New Approach to the Estimation of Health Status Adjusted Life Expectancy,* ***Michael C. Wolfson***

35. *Canadian Jobs and Firm Size: Do Smaller Firms Pay Less?,* ***René Morissette***

36. *Distinguishing Characteristics of Foreign High Technology Acquisitions in Canada's Manufacturing Sector,* ***John R. Baldwin and Paul K. Gorecki***

37. *Industry Efficiency and Plant Turnover in the Canadian Manufacturing Sector,* ***John R. Baldwin***

38. *When the Baby Boom Grows Old: Impacts on Canada's Public Sector,* ***Brian B. Murphy and Michael C. Wolfson***

39. *Trends in the distribution of Employment by Employer Size: Recent Canadian Evidence,* ***Ted Wannell***

40. *Small Communities in Atlantic Canada: Their Industrial Structure and Labour Market conditions in the Early 1980s,* ***Garnett Picot and John Heath***

41. *The Distribution of Federal/Provincial Taxes and Transfers in rural Canada,* ***Brian B. Murphy***

42. *Foreign Multinational Enterprises and Merger Activity in Canada,* ***John Baldwin and Richard Caves***

43. *Repeat Users of the Unemployment Insurance Program,* ***Miles Corak***

44. *POHEM -- A Framework for Understanding and Modelling the Health of Human Population,* ***Michael C. Wolfson***

45. *A Review of Models of Population Health Expectancy: A Micro-Simulation Perspective,* ***Michael C. Wolfson and Kenneth G. Manton***

46. *Career Earnings and Death: A Longitudinal Analysis of Older Canadian Men,* **Michael C. Wolfson, Geoff Rowe, Jane Gentleman and Monica Tomiak**

47. *Longitudinal Patterns in the Duration of Unemployment Insurance Claims in Canada,* **Miles Corak**

48. *The Dynamics of Firm Turnover and the Competitive Process,* **John Baldwin**

49. *Development of Longitudinal Panel Data from Business Registers: Canadian Experience,* **John Baldwin, Richard Dupuy and William Penner**

50. *The Calculation of Health-Adjusted Life Expectancy for a Multi-Attribute Utility Function: A First Attempt,* **J.-M. Berthelot, R. Roberge and M.C. Wolfson**

51. *Testing The Robustness of Entry Barriers,* **J. R. Baldwin, M. Rafiquzzaman**

52. *Canada's Multinationals: Their Characteristics and Determinants,* **Paul K. Gorecki**

53. *The Persistence of unemployment: How Important were Regional Extended Unemployment Insurance Benefits?* **Miles Corak, Stephen Jones**

54. *Cyclical Variation in the Duration of Unemployment Spells,* **Miles Corak**

55. *Permanent Layoffs and Displaced Workers: Cyclical Sensitivity, Concentration, and Experience Following the Layoff,* **Garnett Picot, Wendy Pyper**

56. *The Duration of Unemployment During Boom and Bust*,* **Miles Corak**

57. *Getting a New Job in 1989-90 in Canada,* **René Morissette**

58. *Linking survey and administrative data to study determinants of health,* **P. David, J.-M. Berthelot and C. Mustard**

59. *Extending Historical Comparability in Industrial Classification,* **John S. Crysdale**

60. *What is Happening to Earnings Inequality in Canada?,* **R. Morissette, J. Myles and G. Picot**

For further information, contact the Chairperson, Publications Review Committee, Analytical Studies Branch, R.H. Coats Bldg., 24th Floor, Statistics Canada, Tunney's Pasture, Ottawa, Ontario, K1A 0T6, (613) 951-8213.